DTU Library

At-site and regional modeling of extreme hydrologic events

Madsen, Henrik

Publication date:
1996

Document Version
Publisher's PDF, also known as Version of record

AT-SITE AND REGIONAL MODELLING OF EXTREME HYDROLOGIC EVENTS

Henrik Madsen

Department of Hydrodynamics and Water Resources
Technical University of Denmark

Ph.D. thesis
February 1996

PREFACE

The present thesis is prepared as one of the requirements for the Ph.D. degree. The thesis consists of a summary and four adjoined papers

[1] Madsen, H. and Rosbjerg, D., The partial duration series method in regional index-flood modelling, Water Resources Research, 33(4), 2007, 737-746.
[2] Madsen, H., Rasmussen, P.F. and Rosbjerg, D., Comparison of AMS and PDS methods for modelling extreme hydrologic events. I: At-site modelling, Water Resources Research, 33(4), 2007, 747-757.
[3] Madsen, H., Pearson, C.P. and Rosbjerg, D., Comparison of AMS and PDS methods for modelling extreme hydrologic events. II: Regional modelling, Water Resources Research, 33(4), 2007, 759-767.
[4] Madsen, H. and Rosbjerg, D., Generalized least squares and empirical Bayes estimation in regional PDS index-flood modelling, Water Resources Research, 33(4), 2007, 771-781.

In the text they will be referred to by using [1] – [4], while a usual reference is given to papers included in the reference list.

The major part of the study was carried out at the Department of Hydrodynamics and Water Resources (ISVA) at the Technical University of Denmark (DTU) under the supervision of Associate Professor Dan Rosbjerg. His support and guidance throughout the study is gratefully acknowledged. Half a year was spent at the National Institute of Water and Atmospheric Research (NIWA) in Christchurch, New Zealand. I wish to express my appreciation to Charles Pearson for his support during my stay at NIWA. I also wish to thank Peter Funder Rasmussen, Institut National de la Recherche Scientifique-Eau (INRS-Eau), University of Québec for fruitful cooperation.

Funding of the study was provided by a fellowship from DTU. Financial support for the stay in New Zealand was obtained from the Danish Research Academy.

Lyngby, February 1996

Henrik Madsen

ABSTRACT

At-site and regional estimation of extreme hydrologic events based on the partial duration series (PDS) method has been analysed. The PDS model comprises the assumptions of a Poisson distributed number of threshold exceedances and generalized Pareto distributed exceedance magnitudes, corresponding to a generalized extreme value distribution for the annual maximum series (AMS). At-site T-year event estimation in AMS and PDS has been compared using three different estimation methods, respectively, the maximum likelihood method, the method of moments, and the method of probability weighted moments. In common applications, which correspond to heavy-tailed distributions, the PDS model with the method of moments estimation is generally to be preferred. A regional index-flood method based on PDS data where the regional parameters are estimated using weighted L-moment ratios has been introduced. The performance of the method is evaluated by comparing with at-site estimation and estimation based on the corresponding AMS index-flood method. Even in strongly heterogeneous regions, regional estimation is superior to at-site estimation, and, compared to the AMS procedure, the regional PDS method is more robust with respect to violation of the basic homogeneity assumption of the index-flood method. Procedures to define homogeneous regions and determine regional parent distributions have been discussed and applied to New Zealand flood records. Flood frequency groups defined in terms of catchment characteristics appear more homogeneous with respect to L-moment ratios for PDS than AMS data. Moreover, in determination of the regional parent distribution using L-moment ratio diagrams, PDS data, in contrast to AMS data, provide an unambiguous interpretation. A regional method has been introduced that combines the index-flood model with an empirical Bayes procedure. The prior information of the PDS parameters is inferred from regional data using generalized least squares (GLS) regression that accounts for intersite dependence and regional heterogeneity. In the case of a strongly heterogeneous intersite correlation structure, the GLS procedure provides a more efficient estimate of a regional parameter as compared to the usually applied record-length-weighted average procedure. In addition, the GLS procedure offers a general framework for a reliable assessment of regional homogeneity and parameter uncertainty. A linear Bayes estimation procedure results in reasonable and simple estimates of the T-year event and the associated uncertainty at both gauged and ungauged sites.

RESUMÉ

Madsen, H., 1996, Lokal og regional modellering af ekstreme hydrologiske hændelser, Ph.d. afhandling, Institut for Strømningsmekanik og Vandresurser (ISVA), Danmarks Tekniske Universitet.

Nærværende afhandling omhandler estimation af ekstreme hydrologiske hændelser på såvel lokal som regional skala. Udgangspunktet er en statistisk overskridelsesmodel (dvs. modellering af hændelser over et givet niveau; benævnt en PDS model), der forudsætter at fremkomsten af overskridelser kan beskrives ved en Poisson proces og at overskridelsernes størrelse kan beskrives ved en generaliseret Pareto fordeling. Dette svarer til en generaliseret ekstremværdi fordeling for modellering af årlige maksima (benævnt AMS model). Lokal estimation af *T*-års hændelser baseret på henholdsvis AMS og PDS modellen er sammenlignet ved brug af tre forskellige estimationsmetoder; maksimum likelihood metoden, moment metoden, samt en metode baseret på sandsynlighedsvægtede momenter (benævnt L-moment metoden). I de flest forekommende tilfælde i praksis, svarende til en ekstremværdifordeling med en lang hale (positiv skævhed), giver PDS modellen ved brug af moment metoden generelt de mest pålidelige estimater. En regional estimationsmetode baseret på PDS modellen er introduceret. Modellen forudsætter, at standardiserede data fra forskellige steder i en given region er ensfordelte (homogenitetsantagelse). De regionale model-parametre estimeres ved brug af vægtede middelværdier af de lokale L-moment estimater. Estimation af *T*-års hændelser med den regionale PDS model er sammenlignet med henholdsvis estimation baseret alene på lokale data og estimation baseret på den tilsvarende regionale AMS model. Selv i stærkt heterogene regioner er regional estimation at foretrække frem for lokal estimation. Sammenlignet med den regionale AMS model er PDS modellen mere robust i de tilfælde, hvor den regionale homogenitetsantagelse ikke er opfyldt. Forskellige metoder til at definere homogene regioner og bestemme en regional fordeling er diskuteret og anvendt på et regionalt datasæt af maksimumsafstrømninger fra New Zealand. Grupper af oplande defineret ud fra fysiske oplandskarakteristika afspejler en større homogenitet for PDS data end for AMS data, målt i henhold til variabiliteten af L-momenter i regionen. Ved bestemmelse af den regionale fordelingsfunktion ved brug af L-moment diagrammer giver PDS data en mere entydig fortolkning end AMS data. En regional metode, der inkluderer en bayesiansk beskrivelse af modellens parametre, er introduceret. A priori information i Bayes modellen estimeres ud fra regionale data ved brug af en generaliseret mindste kvadraters metode (benævnt GLS metode), der tager hensyn til

dels afhængighed mellem de enkelte stationer i regionen og dels regional heterogenitet. Såfremt korrelationsstrukturen i regionen er meget heterogen, giver GLS metoden et bedre estimat af en regional modelparameter end den traditionelle brug af en vægtet middelværdi, der er vægtet i henhold til de enkelte dataseriers længde. Desuden giver GLS proceduren en realistisk vurdering af regional homogenitet og parameterusikkerhed. En simpel lineær Bayes metode giver pålidelige estimater af *T*-års hændelsen og den tilhørende usikkerhed for både målte og umålte oplande.

TABLE OF CONTENTS

1. INTRODUCTION

Analysis of extreme hydrologic events is an important element in the design and operation of hydraulic structures. The first stage in the design of reservoir and flood protection structures, such as dams, spillways and dikes, is selection of the design discharge which is determined on the basis of an analysis of extreme streamflows. Estimation of extreme precipitation forms the basis for design of urban drainage systems, and low flow characteristics are important in the design and operation of reservoirs. The design event is usually expressed in statistical terms as the event corresponding to a specified exceedance probability or, equivalently, a specified return period T. Correspondingly, a T-year event is the level which on the average is exceeded once in T years. Ideally, economic analysis should be used to determine optimal designs, i.e. the design corresponding to the point where the total cost of construction and damage is minimum. In most cases, however, the design of a given hydraulic structure is based on regulatory rules. Typical design return periods are in the range 100-1000 years for dikes, 50-100 years for bridges and 2-10 years for storm sewers [*Smith*, 1993]. Spillways for major dams, where the consequences of failure can be quite catastrophic, are usually designed corresponding to the probable maximum flood (PMF) which has no direct frequency interpretation. It is, however, becoming more common to design a primary spillway for much less than the PMF, corresponding to a return period in the range 1000-10,000 years, and then provide an emergency spillway to make up the difference [*Smith*, 1993].

Compared to other elements in the design sequence, including geotechnical, hydraulic and structural design, which can usually be achieved with a relatively high degree of precision, determination of the design event can be subjected to large uncertainties. The uncertainty is generally a combination of sampling and model uncertainty and is especially pronounced when estimating design events corresponding to return periods beyond the observation period (extrapolation). Sampling uncertainty originates from the fact that a limited set of data (a sample) is used to estimate the parameters of a specified statistical distribution, and it is the dominant error source for small return periods. Model uncertainty is relatively more important for extrapolation, but as opposed to sampling uncertainty, which can usually be treated analytically, it is more difficult to handle. It is by no means a simple task to determine the type of frequency distribution to be used in a given situation, and hence the degree of sophistication of the statistical model. Evidently, a thorough understanding of the statistical nature of extreme hydrologic events and the inherent uncertainties is a prerequisite for efficient design of hydraulic structures.

At present, two different methods for extreme value analysis in hydrology are prevalent; the annual maximum series (AMS) method and the partial duration series (PDS) method, also referred to as the peak over threshold (POT) method. Application of the AMS method dates back to the beginning of this century (according to *Kirby and Moss* [1987], *Fuller* [1914] was the first to interpret annual flood flows in terms of probabilities), whereas the PDS method, the subject of the present thesis, is more recent (the formal development of the method is mainly due to *Shane and Lynn* [1964] and *Todorovic and Zelenhasic* [1970]). A fundamental difference between the two methods is the definition of the extreme value region. While the PDS method includes all exceedances above a certain threshold level, the AMS method considers only the annual maximum, notwithstanding that secondary events in a year may exceed annual maxima of other years and that an annual maximum flood in a very dry year hardly can be classified as an extreme flood event. Thus, intuitively the PDS method seems to be the most appropriate method. However, the AMS method is far the most frequently applied method in practice. A survey conducted by the World Meteorological Organization [*WMO*, 1989] comprising 55 agencies in 28 countries revealed that only three agencies recommended the PDS method for flood frequency analysis. Tradition, of course, plays an important role in this respect. However, the main reason for the PDS method being much less applied in practice is probably the lack of an appropriate practical guideline for defining the PDS. The AMS is easy to define, and it is usually reasonable to assume that annual maxima are independent, which is a common presumption in statistical inference. On the other hand, the PDS method involves the choice of an appropriate threshold level, and, in addition, some criteria for selecting the relevant exceedances usually have to be imposed in order to ensure independence between successive peaks. These practical problems are addressed in the present thesis.

Although practical advantages and drawbacks are relevant when choosing between different model candidates, a model comparison should primarily be based on appropriate performance criteria, e.g. the accuracy by which quantiles are estimated. *Cunnane* [1973] compared estimation in the traditional PDS model with exponential distributed threshold exceedances with the corresponding AMS model based on the Gumbel distribution and found that the PDS model is more efficient than the AMS model if the PDS contains more than 1.65 exceedances on average per year. One of the objectives of the present thesis is to generalize Cunnane's result to a larger class of distributions in order to provide specific guidelines for choosing between the AMS and PDS method.

For estimation of extreme events, both sampling and model uncertainty are affected by the amount of data available at the site of interest. To obtain more efficient estimates, additional information should be included. One approach is the use of historical or paleohydrological data. Historical flood information includes, for instance, records of large floods from historical documents and flood markers, whereas paleoflood information is obtained from botanical and geophysical evidence. Another approach, which is the main topic of the present thesis, is the use of regional information. Regionalization is a combination of data from different sites in a region that can be assumed to have similar extreme hydrologic behaviour, i.e. space substitutes time to compensate for a short record at a specific site. The use of regional information reduces the sampling uncertainty by introducing more data, and, in addition, it facilitates the choice of an appropriate statistical distribution. Moreover, regionalization forms the basis for making inferences at ungauged sites, which is extremely important for a general assessment of extreme events.

The regional method used in this study is the so-called index-flood method, which has gained increasing interest in recent years. So far, however, the method has been applied only to AMS. The main objective of the present thesis is to introduce and evaluate the performance of a regional PDS index-flood procedure. In a review of recent advances in flood frequency analysis, *Bobée and Rasmussen* [1995] state that: "Little has been done to include PDS models in a regional estimation scheme, and this is perhaps the main reason that PDS analysis remains less used in practice than the annual flood method. Future research should focus on developing regional estimation procedures for use with PDS data". The present thesis can be seen as a contribution to this research.

In the following are summarized the theoretical concepts and main results of the adjoined papers. It is aimed to put the present study into perspective with previous research; however, a comprehensive state of the art review has not been intended. In Chapter 2, the PDS model is introduced, and basic elements, such as parameter estimation and choice of frequency distribution, are presented with special emphasis given to the recently developed theory of L-moments. Furthermore, at-site quantile estimation in AMS and PDS are compared. In Chapter 3, the regional PDS index-flood model is introduced, and the performance of the model is evaluated with respect to its robustness to violation of the basic assumptions of regional homogeneity and intersite independence. Regional estimation with this model is compared to that of the corresponding AMS index-flood model. Another important element in regional analysis is the grouping of sites into homogeneous regions. This aspect is considered in Chapter 4 together with the closely related topic of determination of a regional parent distribution. In Chapter 5, new estimation procedures for the index-flood method are introduced, including generalized least squares and Bayesian estimation techniques. Finally, a summary and the main conclusions are given in Chapter 6.

2. AT-SITE MODELLING

In this thesis, the classical distinction between at-site and regional analysis is made. At-site modelling employs site specific data only, whereas regional modelling also includes data from other sites in a predefined region of interest. In this chapter, at-site modelling is considered, including a presentation of the applied PDS model, a brief review of different estimation procedures, and a comparative study of T-year event estimation in AMS and PDS, respectively. The following chapters will focus on the regional aspects.

2.1 The PDS model

The PDS model used herein is based on a fixed threshold level, thus being different from the alternative PDS model based on a fixed number of extreme events [*Buishand*, 1989]. First, the definition of PDS related to extreme streamflow analysis is considered. By introducing a threshold level q_0 in a series of streamflows and considering only periods where the streamflow exceeds q_0, a PDS is obtained (see Fig. 1). Thus, if Q_i denotes a peak discharge exceeding q_0, the basic variable to be considered in the following is given by $X_i = Q_i – q_0$. In the case of multiple peaks corresponding to the same event, only the maximum peak of the cluster is considered. To ensure independence between peak flows in the PDS, some restrictions usually have to be imposed on the inter-arrival time of successive peaks and the magnitudes of the interevent discharges. In the United States Water Resources Council guidelines [*USWRC*, 1982] peak floods are considered independent if (1) the interevent time exceeds 5+ln(AREA) days where AREA is the catchment area in square miles, and (2) the interevent discharge drops below 75% of the lowest of the two peaks. A slightly different procedure was proposed by *Cunnane* [1979].

The PDS method has been applied in other hydrologic studies, such as extreme precipitation and streamflow drought analyses, as well as in related disciplines, including ocean and wind engineering. Application examples in precipitation studies include *Van Montfort and Witter* [1986], *Fitzgerald* [1989] and *Madsen et al.* [1994, 1995]. To ensure independence in PDS of daily rainfall extremes, *Fitzgerald* [1989] only included those peaks where no higher rainfall occurred within the preceding or following three days, whereas no specific restrictions were imposed in PDS of extreme rainfall intensities by *Madsen et al.* [1994, 1995]. *Zelenhasic and Salvai* [1987] applied the PDS method in the modelling of streamflow drought duration and deficit volume, and different procedures for pooling mutually dependent droughts were considered by *Madsen and Rosbjerg* [1995] and *Tallaksen et al.* [1997]. *Rosbjerg and Knudsen* [1984] and *Abild et al.* [1992] employed the PDS method for estimating, respectively, significant wave heights and extreme wind speeds. In the following, focus is on flood frequency analysis. However, it should be kept in mind that the methods presented in this thesis have broader possibilities for application.

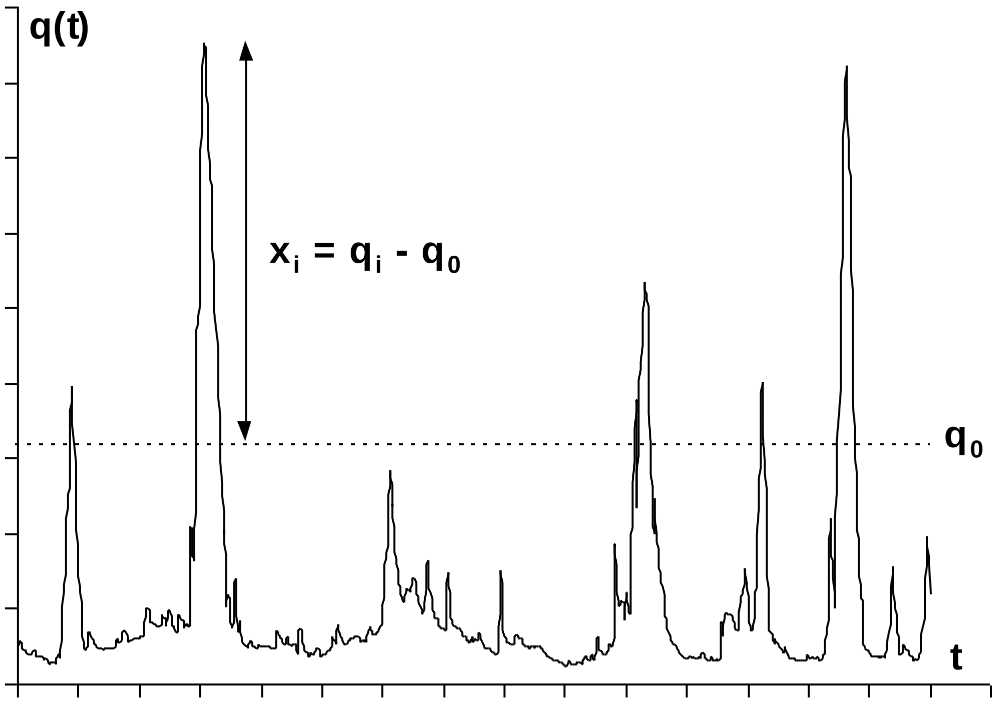

Figure 1 Extraction of peaks from a continuous hydrograph.

Recent reviews of the PDS method are given by *Rasmussen* [1991], *Rosbjerg* [1993] and *Rasmussen et al.* [1994]. The peaks in the PDS are assumed to occur according to a Poisson process. Hence, if the process has an annual periodicity, which seems reasonable in a hydrologic context, the number, N, of exceedances in t years is Poisson distributed with probability function

$$P\{N(t)=n\} = \frac{(\lambda t)^n}{n!} \exp(-\lambda t) \quad , n = 0,1,2,... \tag{2.1}$$

where λ equals the expected number of threshold exceedances per year. In the basic PDS model [*Shane and Lynn*, 1964; *Todorovic and Zelenhasic*, 1970], the exceedance magnitudes are assumed to be independent identically distributed following the exponential (EXP) distribution. The exceedance model used in the present study is the generalized Pareto (GP) distribution which has gained increasing attention in PDS analysis in recent years [e.g. *Van Montfort and Witter*, 1986; *Hosking and Wallis*, 1987; *Fitzgerald*, 1989; *Davison and Smith*, 1990; *Wang*, 1991; *Rosbjerg et al.*, 1992; *Madsen et al.*, 1994, 1995]. The GP distribution has the cumulative distribution function (CDF)

$$F(x) = \begin{cases} 1 - \exp\left(-\dfrac{x}{\alpha}\right) & ,\kappa = 0 \\ 1 - \left(1 - \kappa\dfrac{x}{\alpha}\right)^{\frac{1}{\kappa}} & ,\kappa \neq 0 \end{cases} \tag{2.2}$$

where α and κ are the scale and shape parameter, respectively. For $\kappa = 0$, the EXP distribution is obtained as a special case. For $\kappa < 0$, (2.2) is a reparameterization of the Pareto distribution, and the extension to $\kappa \geq 0$ was given by *Pickands* [1975]. The range of x is $0 \leq x \leq \infty$ for $\kappa \leq 0$, whereas an upper bound exists for $\kappa > 0$: $0 \leq x \leq \alpha/\kappa$. The mean and the variance of the GP distribution read

$$\mu = \mathrm{E}\{\,\mathrm{X}\,\} = \frac{\alpha}{1+\kappa} \quad , \quad \sigma^2 = \mathrm{Var}\{\,\mathrm{X}\,\} = \frac{\alpha^2}{(1+\kappa)^2(1+2\kappa)} \tag{2.3}$$

The *T*-year event is defined as the $(1-1/\lambda T)$-quantile in the distribution of the exceedances [e.g. *Rosbjerg*, 1985], and hence from (2.2) one obtains

$$\mathrm{x_T} = \mathrm{F}^{-1}\left(1 - \frac{1}{\lambda \mathrm{T}}\right) = \begin{cases} \alpha \ln(\lambda \mathrm{T}) & ,\kappa = 0 \\ \dfrac{\alpha}{\kappa}\left[1 - \left(\dfrac{1}{\lambda \mathrm{T}}\right)^{\kappa}\right] & ,\kappa \neq 0 \end{cases} \tag{2.4}$$

An important property of the GP distribution in a PDS context is the so-called "threshold stability". That is, if X is GP distributed and $h > 0$, then $X' = X-h$ given $X > h$ is also GP distributed with the same shape parameter as X (see [2]). Another important property, which motivates the use of the GP distribution in extreme value analysis, is that the annual maximum distribution corresponding to the PDS/ GP model is a generalized extreme value (GEV) distribution. The GEV distribution has the same shape parameter as the parent GP distribution. In particular, the PDS/ EXP model ($\kappa =$

0) corresponds to the Gumbel (EV1) distribution for annual maxima. In [2] the relation between the PDS/ GP and the AMS/ GEV model is described in more detail.

In a study of the Poisson assumption in the PDS model, *Cunnane* [1979] found that it could be verified in only a few of 26 studied catchments in Great Britain. He introduced the negative binomial distribution but it did not seem to offer any satisfactory improvement. In fact, a misspecification of the distribution of the annual number of exceedances is not critical. The most important element in PDS analysis is the modelling of the exceedance magnitudes. Alternative exceedance distributions that have been proposed include the gamma distribution [*Zelenhasic*, 1970], the Weibull distribution [*Miquel*, 1984; *Ekanayake and Cruise*, 1993], and the log-normal distribution [*Rosbjerg et al.*, 1991]. Both the gamma and the Weibull distribution include the EXP distribution as a special case.

When choosing the appropriate exceedance model, both descriptive and predictive abilities should be taken into account [*Cunnane*, 1987]. In general, the model error can be minimised by introducing more parameters; however this interferes with the principle of parsimony. *Rosbjerg et al.* [1992] showed that in the case where κ in the GP distribution is close to zero the EXP distribution yields more efficient T-year event estimators in terms of root mean square error (RMSE). Thus, if no physical evidence suggests a κ-value different from zero, the EXP distribution should be applied in this case. A similar conclusion was obtained by *Lu and Stedinger* [1992b] who compared T-year event estimation in the GEV and EV1 distributions.

The choice of threshold level is a crucial element in PDS analysis. Although important, this aspect has been treated only superficially in the literature. An optimal choice of threshold should ensure as much relevant information as possible to be included in the analysis (i.e. separate the informative extreme events from the remainder of the series) without violating basic statistical assumptions. *Ashkar and Rousselle* [1987] proposed a method that exploits the fact that the mean and the variance of the annual number of exceedances are equal in the case of a Poisson distribution. The threshold level is then determined as the level where the mean-to-variance ratio equals one. Another method, which is related to the GP assumption of exceedance magnitudes, is based on the use of a mean excess plot, also referred to as a mean residual life plot [*Davison and Smith*, 1990; *Naden*, 1992]. This is a plot of the mean of the exceedances against the threshold level. In the case of GP distributed exceedances, then, above a certain threshold level, the plot should follow a straight line with a slope of $-\kappa/(\kappa+1)$. These methods, however, do not provide unambiguous solutions implying a certain degree of subjectivity to be involved in the choice of threshold, and, in addition, in some cases they may not provide any solution at all. Moreover, they are based solely on satisfying certain distribution assumptions without considering the physical properties of extreme events.

A standardised and objective procedure for selecting the threshold level is a prerequisite for application of the PDS model on a regional scale. Such a method should reflect differences in flow regimes between regions. For instance, in catchments dominated by glacier runoff the number of flood peaks is relatively low as compared to catchments dominated by rainstorms (implying different λ-values), and flood peaks in drier regions are more variable than those in wetter regions (implying different κ-values). *Rosbjerg and Madsen* [1992] recommended a method that to some extent reflects differences in flow regimes. This method is based on a predefined frequency factor k, i.e. $q_0 = \mathrm{E}\{Q\}+k\mathrm{S}\{Q\}$ where $\mathrm{E}\{Q\}$ and $\mathrm{S}\{Q\}$ are, respectively, the mean and the standard deviation of the daily data series. Values of k in the range 3-3.5 have been found appropriate [*Rasmussen and Rosbjerg*, 1991; *Madsen et al.*, 1994]. In [3] and [4] a slightly modified method is employed in which the threshold level is defined as a certain quantile of the daily flow duration curve. This method is shown to be reasonably consistent with respect to reflecting differences in extreme value behaviour in the sense that distinct homogeneous regions, which are defined in terms of the κ-parameter, have significantly different λ-parameters.

2.2 Estimation methods

Having determined the family of distributions to apply in a given situation, the next stage in the analysis is the estimation of the parameters of that distribution and subsequently estimation of the T-year event. Several methods exist for parameter estimation of which the method of moments (MOM) and the maximum likelihood (ML) method have been applied most frequently in hydrology. *Hosking* [1990] introduced L-moments which have become popular, not particularly as a tool for at-site parameter estimation, but primarily in regionalisation, considering estimation of regional parameters, delineation of homogeneous regions, and identification of regional parent distributions. In this section the theory of L-moments is briefly reviewed, and in the following chapters application of L-moments in regional studies is elaborated.

L-moments are defined as linear combinations of expected values of order statistics [*Hosking*, 1990]. The first L-moment (λ_1) is the mean value identical to the first conventional moment. The second L-moment (λ_2) is a measure of the scale or dispersion analogous to standard deviation, and L-moments of order three (λ_3) and four (λ_4) are measures of, respectively, skewness and kurtosis. Analogous to product moment ratios, L-moment ratios, L-coefficient of variation (L-C_v), L-skewness and L-kurtosis, can be defined as

$$\begin{aligned} \tau_2 &= \frac{\lambda_2}{\lambda_1} \equiv \text{L-}C_v \\ \tau_3 &= \frac{\lambda_3}{\lambda_2} \equiv \text{L-skewness} \\ \tau_4 &= \frac{\lambda_4}{\lambda_2} \equiv \text{L-kurtosis} \end{aligned} \tag{2.5}$$

L-moments can be written as linear functions of probability weighted moments (PWM), which can be defined as [*Greenwood et al.*, 1979]

$$\beta_r = E\left\{X[F(X)]^r\right\} \quad , r = 0,1,2,... \tag{2.6}$$

where F(.) is the CDF of *X*. The first four L-moments in terms of PWMs read

$$\begin{aligned} \lambda_1 &= \beta_0 \\ \lambda_2 &= 2\beta_1 - \beta_0 \\ \lambda_3 &= 6\beta_2 - 6\beta_1 + \beta_0 \\ \lambda_4 &= 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \end{aligned} \tag{2.7}$$

Thus, procedures based on PWMs and L-moments are equivalent. However, L-moments are more convenient with respect to summarizing a probability distribution. Population L-moments for a number of distributions commonly applied in hydrology are given in *Hosking* [1990] and *Stedinger et al.* [1993].

For estimation of L-moments, *Hosking and Wallis* [1995] recommended the use of unbiased PWM estimators. Given an ordered sample of size *N*, $x_{(N)} \le x_{(N-1)} \le ... \le x_{(1)}$, the unbiased PWM estimators can be written as [*Landwehr et al.*, 1979]

$$b_r = \sum_{i=1}^{N-r} \frac{(N-i)(N-i-1)\cdots(N-i-(r-1))}{N(N-1)(N-2)\cdots(N-r)} x_{(i)} \quad , r = 0,1,2,... \tag{2.8}$$

A library of FORTRAN routines is available for estimation using L-moments (PWM estimation) for a number of different distributions [*Hosking*, 1991]. Compared to product moment ratio estimators, L-moment ratio estimators have smaller bias, and, in addition, they are relatively insensitive to outliers. These features are particularly important in regional studies. In at-site analysis, however, the interest is the estimation of a *T*-year event rather than estimation of particular moments, and, although L-moment estimators have preferable sampling properties, the same does not necessarily apply for the *T*-year event estimator.

Procedures for *T*-year event estimation in the PDS/ GP and AMS/ GEV distributions in the cases of, respectively, ML, MOM and PWM estimation are described in [2]. *Rosbjerg et al.* [1992] compared MOM and PWM estimation for the PDS/ GP model and found that MOM estimation is more efficient (in terms of RMSE of the *T*-year event estimator) except for very large sample sizes, rarely available in practice. *Hosking and Wallis* [1987] employed also the ML method and found that it is preferable to MOM and PWM estimation only for larger sample sizes and for $\kappa > 0.2$. For estimation in the AMS/ GEV model, *Hosking et al.* [1985b] compared ML and PWM estimation and showed that PWM estimation, in general, is more efficient. In [2] MOM estimation in the AMS/ GEV model is considered. Compared to ML and PWM estimation, MOM estimation is generally preferable, except for very small (negative) κ where PWM estimation is more efficient and for very large (positive) κ where ML estimation is more efficient. In summary, these analyses indicate that, if focus is on *T*-year event estimation, MOM estimation is superior in most cases.

In frequency analysis, a quantile estimator should always be accompanied by a statement of its reliability. The uncertainty is usually quantified by using Monte Carlo simulations or by applying asymptotic theory. For the PDS/ GP model, asymptotic expressions for the bias and the variance of the PDS parameter estimators and the *T*-year event estimator were deduced by *Rosbjerg et al.* [1992] in the case of both MOM and PWM estimation. Revised formulae based on Monte Carlo simulations for PWM estimation are given in [1] for application in the case of small samples and small (negative) values of κ. A complete set of asymptotic expressions for the variance of the T-year event estimator for the PDS/ GP and AMS/ GEV models based on, respectively, ML, MOM and PWM estimation is presented in [2].

2.3 Comparison of AMS and PDS at-site T-year event estimation methods

Cunnane [1973] compared estimation in the PDS/ EXP and AMS/ EV1 models using the asymptotic sampling variance of the T-year event estimator as a performance index. For both models, ML estimation was adopted. In [2] the comparative study is reviewed and extended by including also MOM and PWM estimation. In the case of ML estimation, for return periods larger than about 20 years, the PDS/ EXP T-year event estimator has a smaller variance than the AMS/ EV1 estimator if the PDS contains more than 1.64 exceedances on average per year. In the case of MOM estimation, the value of λ in the PDS which yields equal model performance, λ_e, is less than 0.91 for all T (λ_e = 0.91 being the asymptotic value for $T \rightarrow \infty$). For PWM estimation, λ_e is larger than the asymptotic value of 1.24 for all T; specifically for $T < 100$ years, λ_e is larger than 1.64.

The variance of the PDS/ EXP T-year event estimator is the same for all three estimation methods. Thus, for $T > 100$ years, the PDS/ EXP model is preferable provided λ is larger than 1.64, otherwise the AMS/ EV1 model with ML estimation should be employed. For $T < 100$ years, the PDS/ EXP model is preferable provided $\lambda > \lambda_e$ where λ_e is larger than 1.64, e.g. $\lambda_e = 2.86$ for $T = 10$ years and $\lambda_e = 1.85$ for $T = 50$ years, otherwise the AMS/ GEV model with PWM estimation should be used. In conclusion, since λ-values in the range 2-5 are usually obtained in practice, the results indicate that the PDS/ EXP model is preferable in most cases.

In [2] the comparative study is extended to a larger class of distributions by considering T-year event estimation in the PDS/ GP model and the corresponding AMS/ GEV model. The performance of the two models is compared in the cases of, respectively, ML, MOM and PWM estimation using asymptotic theory as well as Monte Carlo simulations. In the case of ML estimation, for all practical purposes, the PDS/ GP model provides the most efficient T-year event estimator irrespective of the values of κ and λ. On the other hand, in the case of MOM estimation, the value of λ in the PDS to obtain equal model performance (λ_e) depends strongly on the κ-parameter.

The PDS/GP model is generally preferable for negative κ, whereas for positive κ, a very large, and in practice unrealistic, number of exceedances is required in the PDS to obtain better performance of the PDS/GP T-year event estimator. Hence, for positive κ the AMS/GEV model is generally to be preferred. In the case of PWM estimation, the results of the model comparison are essentially similar to those obtained for MOM estimation, i.e. the PDS/GP model is generally more efficient for negative κ.

By comparing the performance of the, in total, six different models, a decision rule is formulated that considers the choice of both extreme value model and estimation method. For typical λ-values in the range 2-5, the decision rule is as follows. Generally, for negative κ, the PDS/GP model with MOM estimation should be applied; for $0 < \kappa < 0.2$, the AMS/GEV model with MOM estimation is preferable; and for $\kappa > 0.2$, one should use the PDS/GP model with ML estimation. For small sample sizes, however, the AMS/GEV model with MOM estimation is preferable also for $\kappa > 0.2$. When κ is close to zero, and no physical evidence suggests a κ-parameter different from zero, the PDS/EXP model should be applied. In conclusion, since heavy-tailed distributions, corresponding to negative κ, are far the most common in flood frequency analysis (see e.g. *Farquharson et al.* [1987] and *Gustard et al.* [1989] for comprehensive flood studies) the PDS model is generally to be preferred.

3. REGIONAL MODELLING

It is generally recognized that the most viable way of improving quantile estimation with respect to both descriptive and predictive abilities is regionalisation. Several regional estimation procedures are available (see *Cunnane* [1988] for a comprehensive review). At present, two methods are prevalent; the direct regression method and the index-flood method [*Bobée and Rasmussen*, 1995]. The direct regression method, in which *T*-year events are estimated from physical catchment characteristics using regression techniques, has been widely used in the United States. In the present study the index-flood method is considered. This method has gained increasing interest in recent years; however, applications have been related to AMS only. In this chapter, the index-flood method is reviewed and application of the method for use with PDS data is examined. Moreover, *T*-year event estimation with the AMS and PDS regional index-flood procedures are compared.

3.1 The index-flood method: A review

The index-flood method was originally introduced by *Dalrymple* [1960]. The basic hypothesis of the method is that data at different sites in a region follow the same distribution except for scale. Data in the region are divided by the at-site scale parameter (which serves as the index-flood parameter), and the normalized data are then jointly used to estimate the parameters of the regional distribution. The at-site quantile estimator is subsequently obtained by multiplying the normalized quantile estimator with an estimate of the site specific index-flood parameter. In AMS analysis, the mean annual flood is normally used as the index-flood parameter. Instead of the mean, *Smith* [1989] used a larger quantile.

One approach for estimating the parameters of the regional normalized distribution is the station year method, also referred to as regional pooling of data. In this method the normalized data in the region are treated as if they form a single random sample from the regional distribution, and this sample is then used to estimate the regional parameters. Since the method ignores intersite dependence, it is expected to lead to bias in quantile estimates, especially for large return periods [*Cunnane*, 1988]. *Wallis* [1980] and *Greis and Wood* [1981] introduced an index-flood method in which the parameters of the regional distribution are estimated from regional weighted averages of normalized PWMs or, equivalently, regional weighted average L-moment ratios. By using regional averages, the problem of intersite dependence is less severe in the sense that it does not introduce any bias into quantile estimation, although it does increase the sampling variance [*Stedinger*, 1983; *Hosking and Wallis*, 1988].

Hosking et al. [1985a] compared the PWM index-flood procedure with the regional method recommended in the UK Flood Studies Report [*NERC*, 1975] and found that the index-flood method based on either a GEV or a Wakeby distribution (5-parameter distribution introduced by *Houghton* [1978]) is superior. *Wallis and Wood* [1985] and *Potter and Lettenmaier* [1990] compared the PWM index-flood procedure with the US Water Resources Council method [*USWRC*, 1982] using, respectively, Monte Carlo simulations and a resampling method, and both studies recommended the PWM index-flood method based on a GEV or a Wakeby distribution. In recent years, the GEV/ PWM algorithm has been widely used in regional flood frequency studies [e.g. *Pearson*, 1991a; *Pilon and Adamowski*, 1992].

The homogeneity assumption in the index-flood method prescribes that dimensionless product moments, or equivalently L-moment ratios, of order two and higher (i.e. C_v, skewness, kurtosis etc.) are constant in the region. *Lettenmaier et al.* [1987] studied the effect of regional heterogeneity of the GEV/ PWM index-flood method and showed that the regional estimator performs better than the at-site estimator in moderate heterogeneous regions. In more heterogeneous regions, *Lettenmaier et al.* [1987] recommended the use of a modified index-flood method in which the second moment is estimated from at-site data and only the skewness is based on a regional estimate. *Stedinger and Lu* [1995] also studied this estimator and questioned some of the results by *Lettenmaier et al.* [1987] due to their simulations of unrealistic GEV distributions with large probabilities of negative flows. It should be noted that the modified index-flood method is based on a similar concept as the *USWRC* [1982] method where the skewness in the log Pearson Type 3 distribution is based on a regional estimate. The homogeneity assumption in this case prescribes that dimensionless moments of order three and higher are constant in the region, and hence the method possesses less strict assumptions with respect to regional homogeneity than the usual index-flood method.

Gabriele and Arnell [1991] suggested another variant of the index-flood method based on a hierarchical grouping of sites. Their approach exploits the empirical observation that heterogeneity in flood characteristics is present at different scales, and some characteristics are more variable than others over a given region. In general, the higher order of the regional moment to be estimated, the more stations should be included to obtain a reliable estimate. In the hierarchical GEV/PWM index-flood method, the skewness is estimated from a large region (super-region) and C_v is then estimated from sub-regions within the super-region. *Bobée and Rasmussen* [1995] question the practical feasibility of the hierarchical approach since two sets of homogenous regions have to be identified.

Hosking and Wallis [1988] studied the effect of intersite dependence on the performance of the GEV/PWM index-flood estimator and found that the effect is negligible for moderate correlations. Even when both intersite correlation and modest regional heterogeneity are present, regional estimation is more efficient than at-site estimation.

3.2 The PDS index-flood method

An index-flood method based on the PDS/GP model is introduced in [1]. Using the mean of the exceedances as the index-flood parameter, the regional estimator of the *T*-year exceedance event at site no. *i* is given by

$$\hat{x}_{Ti} = \hat{\mu}_i \, \hat{z}_T \quad , \quad \hat{z}_T = \begin{cases} \ln(\hat{\lambda}_i \, T) & , \kappa^R = 0 \\ \dfrac{1 + \hat{\kappa}^R}{\hat{\kappa}^R} \left[1 - \left(\dfrac{1}{\hat{\lambda}_i \, T} \right)^{\hat{\kappa}^R} \right] & , \kappa^R \neq 0 \end{cases} \tag{3.1}$$

where μ_i is the at-site mean value of the exceedances, λ_i is the at-site Poisson parameter, and z_T is the regional normalized quantile. The PDS/GP index-flood model presumes homogeneity with respect to the shape parameter κ, which is estimated on the basis of weighted averages of normalized PWMs (or L-moment ratios). Usually, the sample size N is used as weight since the sampling uncertainty of an at-site estimator is inversely proportional to N. A different weighting scheme based on generalized least squares estimation is presented in Chapter 5.

Approximate expressions for the bias and the variance of the regional *T*-year event estimator are deduced in [1]. The variance of $\hat{x}_{Ti}$ is approximately given by

$$\mathrm{Var}\{\hat{x}_{Ti}\} = \mu_i^2 \, \mathrm{Var}\{\hat{z}_T\} + z_T^2 \, \mathrm{Var}\{\hat{\mu}_i\} \tag{3.2}$$

First, consider the case of a homogeneous region. As more stations are included in the region, $\mathrm{Var}\{\hat{z}_T\}$ is reduced, and the variance of the regional estimator tends to a lower limit, $z_T^2\mathrm{Var}\{\hat{\mu}_i\}$, that depends only on the number of at-site data. Thus, at least from a predictive point of view, beyond a certain point the gain in including additional stations to the region becomes insignificant. For T = 1000 years, $\mathrm{Var}\{\hat{x}_{Ti}\}$ is virtually constant for a number of stations in the region, M, about 30; for T = 100 years, the constant level is reached for M about 20; and for T = 10 years, only 5 stations are required. In other words, inclusion of regional information is more important when focus is on higher quantile estimation. Essentially similar results were obtained for the AMS/ GEV index-flood model by *Stedinger and Lu* [1995]. When heterogeneity is present, the number of sites required to reach the constant level of the uncertainty of the regional estimator is virtually similar to that obtained in the homogeneous case. However, the uncertainty of $\hat{x}_{Ti}$ is larger than $z_T^2\mathrm{Var}\{\hat{\mu}_i\}$ due to the bias of $\hat{x}_{Ti}$ caused by regional heterogeneity.

The effect of heterogeneity is analysed in more detail in [1] by comparing the performance of the regional estimation procedure with that of estimation based on at-site data only. If heterogeneity of the κ-parameter is present, the regional estimator is still superior to the at-site estimator for small to moderate sample sizes, even in extremely heterogeneous regions, the performance being relatively better in regions with a negative shape parameter. For larger sample sizes, the regional estimator is preferable in homogeneous and moderately heterogeneous regions.

In [1] also the effect of intersite dependence is analysed. It is shown that the effect on the normalized regional quantile estimator $\hat{z}_T$ is well described by *Stedinger*'s [1983] formula. That is, the increase of the variance of the normalized regional T-year event estimator due to correlation depends on the squared correlation coefficient between concurrent exceedances. In other words, in a region of M correlated sites, $\mathrm{Var}\{\hat{z}_T\}$ is the same as for a region of M_E independent sites (denoted the effective number of independent sites) with M_E given by

$$M_E = \frac{M}{1 + (M - 1)\overline{\rho^2}} \tag{3.3}$$

where $\overline{\rho^2}$ is the regional average of the squared correlation coefficient between concurrent exceedances. However, since the uncertainty of $\hat{x}_{Ti}$ is a combination of the uncertainties of, respectively, $\hat{z}_T$ and the estimated at-site mean, cf. (3.2.), the effect of intersite dependence on $\hat{x}_{Ti}$ is much less than predicted by (3.3). For instance, using the symmetrical intersite correlation structure applied in [1] with an average correlation coefficient of 0.4 and M = 20, M_E is about 4, whereas a similar interpretation of the effect of intersite dependence with respect to estimation of x_{Ti} yields an effective number of independent sites of about 14. In the case of both intersite dependence and regional heterogeneity, the analysis shows that the regional estimator is superior to the at-site estimator in modest to strongly heterogenous regions and for moderate correlations. In conclusion, the PDS/ GP regional index-flood method is a robust and efficient estimation method.

A general use of the regional index-flood method involves estimation at ungauged sites, and two aspects have to be considered in this respect. First, the ungauged site has to be assigned a region where the regional normalized frequency curve is known. Procedures for grouping of sites into homogeneous regions are discussed in Chapter 4. Secondly, since no at-site data are available, the index-flood parameter, the Poisson parameter, and the threshold level have to be inferred from regional data. Estimation of the index-flood parameter and the Poisson parameter can be achieved from regression analysis that relates the at-site values to catchment characteristics, such as geologic, physiographic and meteorologic characteristics. Regression analysis is discussed in Chapter 5. For estimation of the threshold level, if expressed in terms of a given quantile of the daily flow duration curve as proposed above, the method given by *Fennessey and Vogel* [1990] where the flow duration curve is estimated from two characteristics, catchment area and a basin relief parameter, can be used.

3.3 Comparison of AMS and PDS regional index-flood estimation methods

The PDS/ GP and AMS/ GEV regional index-flood procedures are compared in [3] with respect to estimation in both homogeneous and heterogeneous regions. Regional heterogeneity is expressed in terms of the regional variability of κ. To simulate AMS/ GEV data corresponding to the parent PDS/ GP model, a realistic relationship between C_v and κ in regional GEV distributions is employed [*Farquharson et al.*, 1987; *Lu and Stedinger*, 1992b]. Hence, the regional variability of C_v of AMS is implicitly expressed in terms of the regional variability of κ.

For estimation in homogeneous regions, virtually similar results are obtained as in the case of at-site estimation using PWM estimation. The value of λ in the PDS to obtain equal model performance (λ_e) increases for increasing κ. For λ-values in the range 2-5, the PDS/ GP index-flood method is generally preferable in regions with a negative κ, whereas the AMS/ GEV index-flood method is more efficient in regions with a positive κ.

For estimation in heterogeneous regions, the PDS/ GP method is relatively more efficient, i.e. as the degree of heterogeneity increases so does the relative efficiency of the PDS/ GP method, and in regions with a realistic degree of heterogeneity the PDS/ GP method is superior. Thus, the PDS/ GP method is more robust than the AMS/ GEV method with respect to violation of the basic homogeneity assumption; a property which is extremely important since real-world regions always possess some degree of heterogeneity. In more heterogeneous regions, the modified AMS/ GEV index-flood method in which only the skewness is estimated from regional data becomes competitive, especially in regions with positive κ-values. In such cases, however, to obtain more efficient T-year event estimates, one should rather try to divide the region into smaller and hence more homogeneous groups than to apply the modified AMS/ GEV index-flood procedure. In conclusion, for estimation in typical regions (corresponding to negative κ-values) with a realistic degree of heterogeneity, the PDS/ GP index-flood procedure is to be preferred.

4. DELINEATION OF HOMOGENEOUS REGIONS

An important aspect in regional analysis concerns the delineation of homogenous regions. Since the homogeneity assumption of the index-flood method prescribes data at different sites in the region to be identical except for scale, identification of a homogeneous region is closely related to the determination of a regional common distribution. In this chapter, procedures for grouping of sites into homogeneous regions and determination of regional parent distributions are discussed.

4.1 Grouping of sites

Traditionally, geographically coherent regions have been applied in regional flood frequency analysis. Geographical regions are convenient for administrative reasons and practical applications; however, geographical proximity does not necessarily imply hydrologic similarity. *Wiltshire* [1986b] analysed the 10 geographical flood frequency regions in Great Britain [*NERC*, 1975] and found that these regions exhibit a significant variability and must be interpreted as heterogeneous. *Matalas et al.* [1975] examined the properties of sample skewness of annual floods in 14 geographical regions in the United States and found that the variability in skewness among the different series in each region was larger than that calculated from data sets generated from various distributions. They referred to this phenomenon as the condition of separation of skewness. The separation effect has been used as an important criterion for choosing flood frequency distributions; that is, for a distribution to be adequate for flood frequency analysis it must reproduce as much variability in skewness as is observed in flood data sets [e.g. *WMO*, 1989]. However, the actual cause of the separation effect has been shown to be due to spatial mixing of different skewness values [*Ashkar et al.*, 1992], i.e. the 14 geographical regions applied by *Matalas et al.* [1975] were in fact heterogeneous. Thus, the separation effect should not be used as a criterion for justifying a specific type of flood frequency distribution.

Recent research has focused on methods that define hydrologic similarity between sites in a multi-dimensional space of flood statistics and/or physical catchment characteristics. Since homogeneity is defined in terms of statistical measures, flood statistics are required for a correct grouping of sites into homogenous regions, whereas inclusion of catchment characteristics is important for a physical understanding of hydrologic similarity. Evidently, assignment of an ungauged site to a region requires the use of physical catchment characteristics.

A widely used regionalization technique is cluster analysis. *Wiltshire* [1986b] applied cluster analysis to define groups in a two-dimensional space of C_v and specific mean annual flood. Subsequently, discriminant analysis was employed to relate these groups to catchment characteristics. While this method virtually ensures a homogeneous grouping of sites, the efficiency of the discriminant analysis is limited, reflecting the, in general, great difficulties of relating flood statistics to catchment characteristics. *Acreman and Sinclair* [1986] used cluster analysis on the basis of catchment characteristics. Their method is more convenient for allocating ungauged sites to predefined clusters, but homogeneity may be more difficult to attain. *Nathan and McMahon* [1990] considered various problems associated with cluster analysis with special emphasis given to the important issue of selecting and weighting variables used to assess similarity between sites. Other application examples of cluster analysis include *Burn* [1989] and *Gustard et al.* [1989].

Wiltshire [1985] proposed a split-sample regionalization approach for defining groups according to catchment characteristics. At its simplest, the method splits a set of basins into two groups based on a single partitioning value of one chosen catchment characteristic. For instance, the basins can be divided into wet and dry basin groups according to average annual rainfall. Measures of variability in each group are evaluated and aggregated into one statistic, and the optimum grouping is then achieved at the partitioning point where the variability statistic is minimum. This process is repeated for other characteristics as well as for a multiple partitioning, i.e. a four-way grouping based on two characteristics, an eight-way grouping based on three characteristics, etc. To measure flood frequency variability, *Wiltshire* [1985] used different statistics based on fitting the GEV distribution to each group, while *Pearson* [1991b], more generally, used a function of L-moment ratios. *Wiltshire* [1986c] employed two statistics that measure, respectively, the variability within groups (which should be minimised) and the variability between groups (which should be maximised).

Delineation of regions in a multi-dimensional space of different characteristics introduces boundary discontinuities, i.e. basins that are hydrologically very similar may fall on either side of the divide between two regions. To avoid abrupt changes in flood quantiles across region boundaries, *Wiltshire* [1986c] introduced the concept of fractional membership of a basin to more than one region. In this case the normalised quantile at the specified site is estimated as a weighted average of the different regional normalised quantiles. *Acreman and Wiltshire* [1989] expanded this idea and proposed a method that dispenses with fixed regions entirely and assigns each site its own region, consisting of those sites that are sufficiently similar to the site of interest. This method was further developed by *Burn* [1990a, 1990b] who referred to the method as a region of influence (ROI) approach. To define the set of sites to be included in the ROI for a given site, a similarity measure is employed based on a weighted Euclidean distance in the multi-dimensional space of selected characteristics between the specified site and the other sites in the region. *Zrinji and Burn* [1994] considered the ROI approach for application at ungauged sites where only catchment characteristics are used to define the distance metric.

To assist in defining homogeneous regions, statistical tests are usually employed. In a homogeneous region all sites have identical population parameters except for scale, and hence a proper homogeneity measure is based on a comparison of the variability between sites of a certain statistical measure and the expected variability of that measure in a homogeneous region (i.e. where the variability is caused by sampling uncertainty only). *Hosking and Wallis* [1993] proposed a test based on L-moment ratios where the expected variability is obtained from Monte Carlo simulations. Denote by V_1 the record-length-weighted standard deviation of the at-site L-C_v estimates. The test statistic reads

$$H = \frac{V_1 - \mu_V}{\sigma_V} \tag{4.1}$$

where μ_V and σ_V are, respectively, the mean and the standard deviation of V_1 obtained from simulations of a homogeneous region with parent L-moment ratios equal to the record-length-weighted averages and record lengths identical to the historical records. In the simulations a four-parameter kappa distribution [*Hosking*, 1988] is adopted. Based on simulation experiments, *Hosking and Wallis* [1993] concluded that a region can be regarded as acceptably homogeneous if $H < 1$; possibly heterogeneous if $1 \leq H < 2$; and definitely heterogeneous if $H \geq 2$. *Hosking and Wallis* [1993] also constructed alternative test statistics based on either L-C_v and L-skewness or L-skewness and L-kurtosis.

Several other homogeneity tests have been proposed based on different variability measures, including C_v [*Wiltshire*, 1986a], non-exceedance probabilities [*Wiltshire*, 1986a], likelihood ratios [*Acreman and Sinclair*, 1986], L-C_v and L-skewness [*Chowdhury et al.*, 1991], and normalised 10-year events [*Lu and Stedinger*, 1992a]. It should be noted that, in general, a homogeneity test statistic has only moderate power for discriminating between homogeneous and heterogeneous regions [e.g. *Wiltshire*, 1986a; *Lu and Stedinger*, 1992a], and hence it should not be used as a strict significance test but rather as a guideline. Also note that the tests usually assume independence between sites. In the case of intersite dependence, the tests are conservative in terms of Type 1 error and their power is reduced [*Lu and Stedinger*, 1992a]. Implications of intersite dependence will be discussed in the following chapter.

In [3] the AMS and PDS regional schemes are applied to flood records from 48 New Zealand catchments. To identify homogenous groupings of sites, the split-sample regionalisation approach is adopted based on two catchment characteristics, average annual rainfall (AAR) and average catchment slope (S). The flood frequency variability measure is based on L-moment ratios where L-C_v is weighted ahead of L-skewness, which in turn is weighted ahead of L-kurtosis, i.e. homogeneity is mainly influenced by L-C_v and less so by L-skewness and L-kurtosis. To test homogeneity of the defined groups, the homogeneity test given by *Hosking and Wallis* [1993] is employed.

For both AMS and PDS data, the regionalization procedure provides well defined groupings for AAR and S considered individually (in a two-way grouping) or combined (in a four-way grouping). However, the defined groups are more homogeneous for PDS than AMS data. For PDS a two-way grouping based on AAR is sufficient to attain homogeneity, whereas for AMS not even a four-way partitioning provides a satisfactory grouping. Thus, for AMS data a further partitioning is necessary, which has the disadvantage of less sites per group and hence larger uncertainties of regional quantile estimates.

4.2 Determination of the regional parent distribution

Having determined a satisfactory grouping of sites, the next stage in regional analysis is the choice of an appropriate statistical distribution to be fitted to the regional data. Due to the small samples usually encountered in hydrology, traditional goodness-of-fit tests that are based solely on at-site data have little power for discriminating between various distributions, and hence procedures for identifying the parent distribution at a regional scale have been advocated by several researchers. In this respect, L-moment ratio diagrams have shown to be a valuable tool [e.g. *Vogel and Fennessay*, 1993].

Table 1 Coefficients of polynomial approximation of L-C_v (τ_2) as a function of L-skewness (τ_3), $\tau_2 = \Sigma A_i \tau_3^i$ (from *Vogel and Wilson* [1996]).

A_i	GP	LN	GAM	WEI
A_0	0.33299	-	-	0.17864
A_1	0.44559	1.16008	1.74139	1.02381
A_2	0.16641	-0.05325	-	-0.17878
A_3	-	-	-2.59736	-
A_4	-	-0.10501	2.09911	-0.00894
A_5	0.09111	-	-	-
A_6	-	-0.00103	-0.35948	-0.01443
A_7	-0.03625	-	-	-

Table 2 Coefficients of polynomial approximation of L-kurtosis (τ_4) as a function of L-skewness (τ_3), $\tau_4 = \Sigma A_i \tau_3^i$ (from *Hosking* [1991]).

A_i	GP	LN	GAM	WEI
A_0	-	0.12282	0.12240	0.10701
A_1	0.20196	-	-	-0.11090
A_2	0.95924	0.77518	0.30115	0.84838
A_3	-0.20096	-	-	0.06669
A_4	0.04061	0.12279	0.95812	0.00567
A_5	-	-	-	0.04208
A_6	-	-0.13638	-0.57488	0.03763
A_7	-	-	-	-
A_8	-	0.11368	0.19383	-

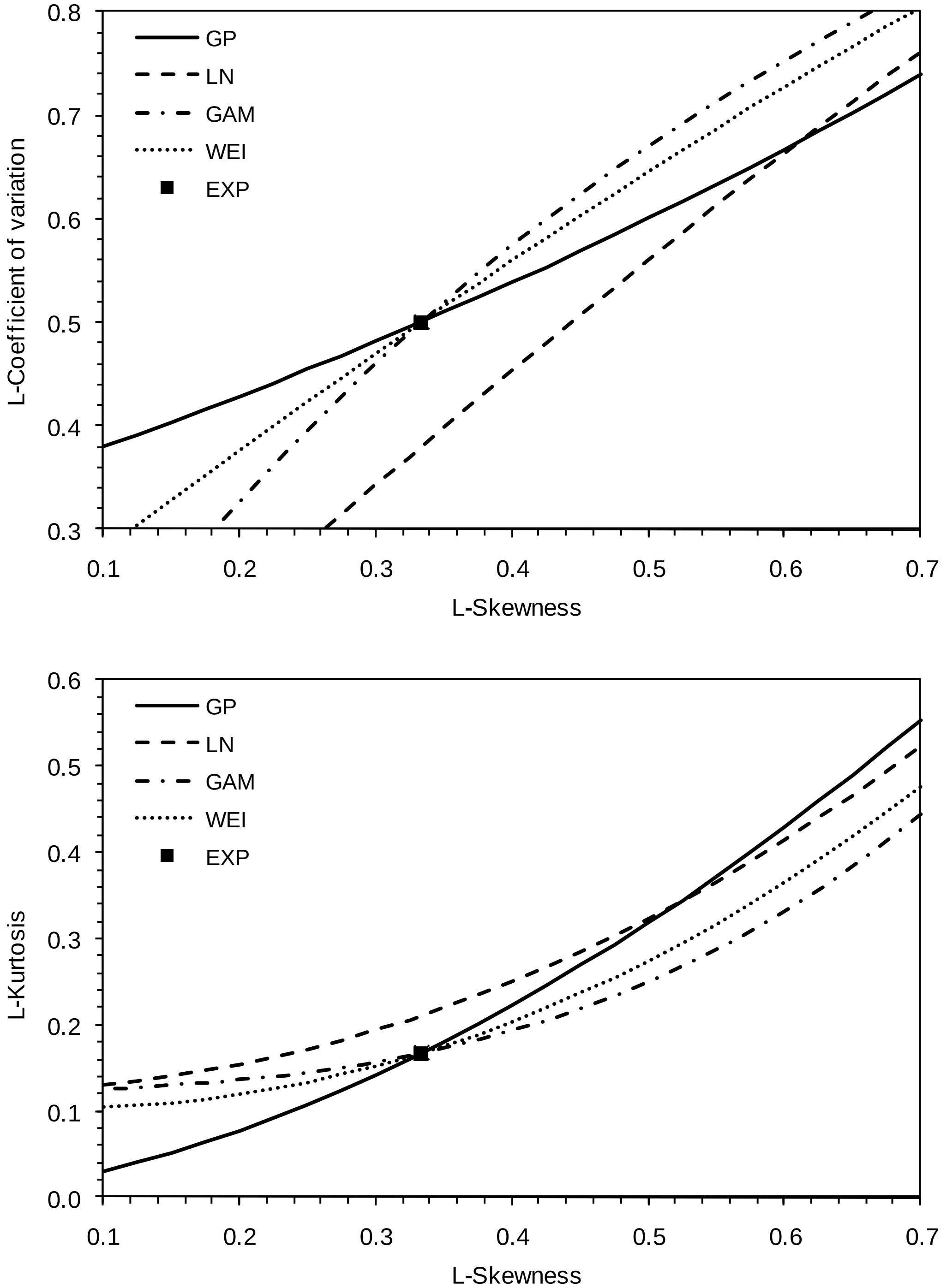

Figure 2 L-moment ratio diagrams, L-C_v versus L-skewness and L-skewness versus L-kurtosis, for the generalised Pareto (GP), log-normal (LN), gamma (GAM), Weibull (WEI), and exponential (EXP) distributions.

L-moment ratio relationships, i.e. L-C_v versus L-skewness and L-skewness versus L-kurtosis, for a number of different distributions that have been proposed for modelling exceedances in a PDS are shown in Fig. 2. These include the GP, the log-normal (LN), the gamma (GAM), the Weibull (WEI), and the EXP distribution (note that the GP, GAM and WEI distributions all include the EXP distribution as a special case). The population L-moments for these distributions are given by *Hosking* [1990] and *Stedinger et al.* [1993]. *Hosking* [1991] and *Vogel and Wilson* [1996] gave polynomial approximations to the L-moment relationships which provide sufficient accuracy in most applications and are easier to apply. The coefficients of the polynomial approximations are given in Table 1 and Table 2.

For a visually based choice of an appropriate distribution, sample estimates of L-C_v, L-skewness and L-kurtosis are in L-moment ratio diagrams compared with the theoretical relationships for a number of parent distributions. Since L-moment ratio estimators are nearly unbiased, approximately half of the at-site sample points in the L-moment ratio diagram are expected to lie above the theoretical curve and half to lie below. To discriminate between various three-parameter distributional alternatives, which is common in AMS analysis, the L-skewness/ L-kurtosis diagram is used. Note that the L-skewness/ L-kurtosis relationship for the different two-parameter distributions in Fig. 2 and Table 2 equals that of their three-parameter counterparts (the three-parameter GAM distribution is often referred to as the Pearson Type 3 distribution). The L-skewness and L-kurtosis for the GEV distribution equal $-\tau_3$ and τ_4 for the WEI distribution. To discriminate between two-parameter distributional alternatives in PDS analysis, the L-C_v/ L-skewness diagram is sufficient.

L-moment ratio diagrams have been used in several regional studies to assess the goodness-of-fit of various three-parameter distributions for modelling annual floods. Studies that consider relatively large regions include: New Zealand, 275 sites [*Pearson*, 1991a]; Australia, 61 sites [*Vogel et al.*, 1993a]; Southwestern USA, 383 sites [*Vogel et al.*, 1993b]; and Continental USA, 1490 sites [*Vogel and Wilson*, 1996]. Interestingly, all these studies (and several other studies from minor regions reported by [*Vogel and Wilson*, 1996]) recommend the use of the GEV distribution. It should be noted, however, that these studies, especially for the large regions, are based on analysing heterogeneous regions, and this heterogeneity may in fact be caused by a mixing of several distinct homogeneous regions with different parent distributions.

To supplement the visual judgement, a goodness-of-fit test may be conducted. A test that is directly related to the L-moment ratio diagram was proposed by *Hosking and Wallis* [1993]. The test is based on the difference between the regional average L-kurtosis $\hat{\tau}_4^R$ and the L-kurtosis of the fitted regional distribution τ_4. The test statistic read

$$Z = \frac{\tau_4 - \hat{\tau}_4^R + \beta_4}{\sigma_4} \tag{4.2}$$

where β_4 and σ_4 are, respectively, the bias and the standard deviation of the regional average L-kurtosis obtained from simulations of a kappa population (similar to calculation of the homogeneity statistic, cf. (4.1.)). The fit is adequate if $|Z| \leq 1.64$, corresponding to an approximate 90% confidence level for accepting the hypothesized distribution. Note that the test is designed for three-parameter distributions to be used in AMS analysis. For a two-parameter candidate distribution in PDS analysis, the test should rather be based on the L-skewness.

In [3] L-moment ratio diagrams are constructed for the AMS and PDS flood data from the 48 New Zealand catchments. For the PDS data, the two-way grouping based on AAR provides virtually two distinct groups of points in the L-skewness/ L-kurtosis space corresponding to each region. The goodness-of-fit test given by *Hosking and Wallis* [1993] reveals that the GP distribution is adequate in both regions. For the AMS data, the points in the L-moment ratio diagram are more dispersed, which makes an unambiguous interpretation more difficult. Both the three-parameter log-normal distribution and the GEV distribution are found to be adequate. In conclusion, the New Zealand application example reveals that with respect to both identification of homogeneous regions and determination of regional distributions the PDS approach has preferable properties.

5. GLS AND EMPIRICAL BAYES ESTIMATION METHODS

Although reasonable care is taken in selecting the sites to be included in the region and statistical tests may justify that C_v and higher order moments are identical in the region, in practice, the homogeneity assumption is never rigorously fulfilled. Moreover, to obtain a region of a reasonable size, one may in some cases be obliged to accept a certain degree of heterogeneity. As shown in Chapter 3, the PDS/ GP index-flood model has shown to be robust with respect to violation of the basic homogeneity assumption, and regional estimation is preferable to at-site estimation even in strongly heterogeneous regions. However, Bayesian methods can be employed for a realistic description of the regional variability and hence provide more efficient regional *T*-year event estimators.

In a Bayesian context, model parameters are treated as stochastic variables, and beliefs or knowledge about the parameters is expressed in terms of probability distributions, referred to as prior distributions. In regional empirical Bayes analysis, the parameters of the prior distribution are obtained from empirical data, generally expressed in terms of catchment characteristics [*Kuczera*, 1982]. Specifically, for the PDS/ GP index-flood model, this approach allows the variability of the regional parameter κ to be modelled. In addition, prior information of the index-flood parameter μ and the Posisson parameter λ can be taken into account. In this chapter an empirical Bayes index-flood estimator is introduced where the prior information is inferred from regional data using generalized least squares (GLS) regression.

5.1 GLS regression

Denote by $\hat{\theta}_i$ a PDS parameter estimator at station no. *i* (i.e. $\hat{\mu}_i$, $\hat{\lambda}_i$ or $\hat{\kappa}_i$). To describe the prior information of θ_i the following linear model is assumed

$$\hat{\theta}_i = \beta_0 + \sum_{k=1}^{p} \beta_k A_{ik} + \varepsilon_i + \delta_i \qquad , i = 1,2,..,M \tag{5.1}$$

or, equivalently, a log-linear relationship. In (5.1), A_{ik}, k = 1,2,..,p are the considered catchment characteristics, ε_i is a random sampling error with $E\{\varepsilon_i\} = 0$, and δ_i is an error term owing to lack of fit of the regression model (model error) with E{δ_i} = 0. The covariance structure of the total errors $\eta_i = \varepsilon_i + \delta_i$ reads

$$\mathrm{Cov}\ \{\eta_{\mathrm{i}}, \eta_{\mathrm{j}}\} = \begin{cases} \sigma^2_{\varepsilon \mathrm{i}} + \sigma^2_{\delta} & , \mathrm{i} = \mathrm{j} \\ \sigma_{\varepsilon \mathrm{i}}\, \sigma_{\varepsilon \mathrm{j}}\, \rho_{\varepsilon \mathrm{ij}} & , \mathrm{i} \neq \mathrm{j} \end{cases} \tag{5.2}$$

where $\sigma_{\varepsilon_i}^{2}$ is the sampling error variance of $\hat{\theta}_i$, σ_{δ}^{2} is the model error variance, and $\rho_{\varepsilon_{ij}}$ is the intersite correlation coefficient. When the residuals are heteroscedastic ($\sigma_{\varepsilon_i}^{2}$ depends on *i*) and cross-correlated ($\rho_{\varepsilon_{ij}} \neq 0$), the GLS method provides better model parameter estimates than the ordinary least squares regression procedure [*Stedinger and Tasker*, 1985], and, in addition, it produces a reasonable and nearly unbiased estimate of the model error variance [*Stedinger and Tasker*, 1986]. For $\rho_{\varepsilon_{ij}}$ = 0, the GLS method corresponds to a weighted least squares (WLS) approach [*Tasker*, 1980]. The GLS regression procedure for estimation of the prior mean and variance of the PDS parameters is described in [4].

An important special case of the regression model, a regional mean model, is obtained when only β_0 is included in (5.1). In the case of homogeneous sampling errors, i.e. identical sampling error variances and intersite correlation coefficients in the region, the prior mean equals the simple regional average of $\hat{\theta}_i$. In general, however, the prior mean is a weighted average of $\hat{\theta}_i$, weighted according to the covariance matrix of the errors, cf. (5.2). In the WLS case ($\rho_{\varepsilon_{ij}}$ = 0) assuming regional homogeneity (σ_{δ}^{2} = 0), the estimate of the prior mean reduces to the record-length-weighted average usually adopted in index-flood modelling. Thus, GLS regression is a general method for inferring regional information where the record-length-weighted average is a special case that considers neither regional heterogeneity nor intersite dependence. GLS regression also provides a consistent uncertainty measure of the regional estimator and an estimate of the model error variance. The latter may be interpreted as a heterogeneity measure; that is, if $\hat{\sigma}_{\delta}^{2}$ > 0, the hypothesis of regional homogeneity may be questioned.

In [4] the regression procedure is applied to the PDS flood records from the two New Zealand regions (denoted Region A and Region B in the following) defined with respect to the average annual rainfall, cf. Chapter 4. Three different regression models are considered (1) WLS regression, (2) GLS regression assuming a homogeneous correlation structure, and (3) GLS regression where the heterogeneity of the correlation structure is taken into account by relating the intersite correlation coefficient to the distance between sites.

With respect to estimation of κ, a regional mean model is adopted. In the case of a strongly heterogenous correlation structure, as observed in Region A, the WLS and GLS procedures are found to differ significantly. In this case erroneous results are obtained if the WLS procedure (corresponding to a record-length-weighted average procedure) or the GLS procedure assuming a homogeneous correlation structure is applied. Note that the analysis of the effect of intersite dependence in [1] did not consider strongly heterogeneous correlation structures as observed in this case. To deal with such a problem, GLS regression should be applied. The results from the WLS regression indicate that both regions are homogeneous ($\hat{\sigma}_\delta^2$ = 0) in agreement with the homogeneity test given by *Hosking and Wallis* (1993). However, in Region B, when intersite dependence is taken into account, the regression results indicate that the region is in fact heterogenous ($\hat{\sigma}_\delta^2$ > 0). Thus, the lack of power of traditional homogeneity tests in the case of intersite dependence may lead to erroneous conclusions with respect to regional homogeneity. Moreover, neglecting intersite dependence and the resulting heterogeneity, may imply a serious underestimation of the uncertainty of the regional κ-estimator. In conclusion, the application example reveals that the GLS procedure provides a more efficient regional estimator than the usually applied record-length-weighted average procedure, and, in addition, it provides a general framework for a reliable assessment of the uncertainty of the regional estimator as well as for an objective appraisal of regional homogeneity.

The regression procedure is also applied to the index-flood parameter μ where a log-linear model is assumed. In both regions, catchment area is found to explain a large part of the regional variability followed by average annual rainfall and, in Region A only, catchment slope. The three estimation methods yield virtually identical results, i.e. intersite dependence is not important in this case. Contradictory to previous regression analyses of mean annual floods [e.g. *Hebson and Cunnane*, 1987], this study indicates a more significant correlation between index-flood parameter and catchment characteristics. The information content of the regional data corresponds to an at-site record length of 4-5 years.

For medium and large quantile estimation (corresponding to T > 10 years), the variability of the λ-parameter can be neglected, cf. [1], and hence a regional mean model for the λ-parameter is sufficient. In the application example, the GLS procedure yields a slightly larger regional estimate of λ than the WLS procedure.

5.2 Bayesian T-year event estimator

The prior mean and variance of the PDS parameters obtained from GLS regression form the basis for inclusion of regional information in empirical Bayes analysis. To make inferences at ungauged sites, a prior *T*-year event estimate is determined on the basis of the prior moments of the PDS parameters. At gauged sites, the prior information is combined with sample information to obtain a posterior *T*-year event estimate. In [4] two different estimators based on the PDS/ GP model are considered, respectively, a linear Bayes estimator that requires only the prior mean and variance of the PDS parameters to be specified, and a parametric Bayes estimator in which the family of prior distributions have to be defined.

In the linear Bayes model, the prior *T*-year event estimator and the associated variance are simply obtained by inserting the prior moments of the PDS parameters in the expressions for the index-flood estimator, cf. (3.1), and its variance, cf. (3.2) (note that the variance of the normalized quantile estimate can be expressed in terms of the prior variances of κ and λ). The posterior mean of a PDS parameter, $\hat{\theta}^{EB}$, is obtained by combining the prior moments ($\hat{\theta}^{REG}$ and Var{$\hat{\theta}^{REG}$}) and the sample information (quantified by $\hat{\theta}^{AS}$ and Var{$\hat{\theta}^{AS}$}) as

$$\hat{\theta}^{\mathrm{EB}} = \mathrm{s}\,\hat{\theta}^{\mathrm{REG}} + (1 - \mathrm{s})\,\hat{\theta}^{\mathrm{AS}} \quad , \quad \mathrm{s} = \frac{\mathrm{Var}\{\hat{\theta}^{\mathrm{AS}}\}}{\mathrm{Var}\{\hat{\theta}^{\mathrm{AS}}\} + \mathrm{Var}\{\hat{\theta}^{\mathrm{REG}}\}} \tag{5.3}$$

with variance Var{$\hat{\theta}^{EB}$} = sVar{$\hat{\theta}^{REG}$}. The factor s expresses the relative weight assigned to, respectively, regional and at-site information, depending on the uncertainty of the two information sources. Under quadratic loss, (5.3) is the best linear estimator irrespective of the type of prior and sample distribution [*Kuczera*, 1983]. The posterior *T*-year event estimator and the associated variance are obtained from (3.1)-(3.2).

The parametric Bayes PDS/ GP model was introduced by *Madsen et al.* [1994] as an extension of a Bayesian model based on the EXP distribution considered by *Rousselle* *and Hindie* [1976] and *Rasmussen and Rosbjerg* [1991]. The model was applied to extreme precipitations by *Madsen et al.* [1994, 1995] and *Madsen and Rosbjerg* [1994], and *Rosbjerg and Madsen* [1996] analysed the efficiency of the model with respect to regional heterogeneity and intersite dependence. The family of prior distributions are, respectively, an inverse gamma distribution for μ, $f_\mu(m)$, a beta distribution for κ, $f_\kappa(k)$, and a gamma distribution for λ, $f_\lambda(l)$. The parameters in these distributions are estimated on the basis of prior moments of the PDS parameters. The prior T-year event distribution is determined by a change of variables

$$f_{x_T}(x) = \int_{-1/2}^{1/2} \int_0^\infty f_\mu(m)\, f_\kappa(k)\, f_\lambda(l) \left| \frac{dg}{dx} \right| \Bigg|_{m=g(x)} dl\, dk \tag{5.4}$$

where the transformation $m = g(x) = kx / [(1+k)(1-[lT]^{-k})]$ is obtained from (2.4). The integral in (5.4) has to be solved numerically. Posterior distributions of the model parameters are obtained by combining prior and site specific information (quantified by the sample likelihood functions) using Bayes' theorem, and subsequently the posterior T-year event distribution is determined from (5.4) by inserting the posterior probability density functions. A point estimator of the T-year event and the associated variance can be determined as, respectively, the mean and the variance in the x_T-distribution.

In [4] the two Bayesian estimation procedures are applied to the New Zealand flood data. The two estimators are found to differ slightly. The main reason for this difference is that the prior distributions of the PDS parameters have positive skewness which is not accounted for in the linear Bayes model. In general, however, the difference between the two estimators is smaller than the uncertainty of the T-year event estimate. Thus, for the sake of computational simplicity, the linear Bayes estimator is sufficient in most cases. If the complete probability distribution of x_T is needed, which may be relevant in more detailed Bayesian studies including, for instance, economical aspects, the parametric Bayes procedure should be applied.

6. SUMMARY AND CONCLUSIONS

At-site and regional estimation of extreme hydrologic events based on the PDS approach have been analysed. The PDS model used in this study comprises the assumptions of a Poisson distributed number of threshold exceedances and GP distributed exceedance magnitudes, corresponding to a GEV distribution for annual maxima with the same shape parameter as in the GP distribution. When the shape parameter equals zero, the PDS exceedance model reduces to the EXP distribution, corresponding to an EV1 distribution for annual maxima.

At-site T-year event estimation in AMS and PDS has been compared in the cases of, respectively, ML, MOM and PWM estimation. For typical λ-values in the range 2-5, the following conclusions were obtained. In the case of ML estimation, the PDS model provides the most efficient T-year event estimator. For MOM and PWM estimation, preference of either model depends strongly on the κ-parameter. The PDS model is generally more efficient for negative κ, whereas the AMS model is preferable for positive κ. A comparison of the six different models, considering the choice of both extreme value model and estimation method, revealed that in general one should use the PDS model with MOM estimation for negative κ; the AMS model with MOM estimation for moderately positive κ; and the PDS model with ML estimation for large positive κ. When κ is close to zero, and no physical evidence suggests a κ-value different from zero, the PDS model assuming EXP distributed exceedances is preferable. Thus, since heavy-tailed distributions, which correspond to negative κ, are far the most common in hydrology, the results obtained in this study suggest that the PDS model in general is to be preferred.

A regional index-flood method based on the PDS model has been introduced. The model presumes homogeneity with respect to the shape parameter, and the mean value serves as the site specific index-flood parameter. The regional shape parameter is estimated on the basis of weighted regional average L-moment ratios. Application of the PDS model on a regional scale requires a standardised and objective procedure for selecting the threshold level. A method that defines the threshold as a certain quantile of the daily flow duration curve has shown to be reasonably consistent with respect to reflecting differences in extreme value behaviour between regions.

The robustness of the regional method with respect to violation of the basic assumptions of homogeneity and intersite independence has been evaluated by comparing the efficiency of the regional estimation procedure with that based on at-site data only. If heterogeneity of the κ-parameter is present, the regional estimator is more efficient than the at-site estimator for small to moderate sample sizes even in extremely heterogeneous regions. For larger sample sizes, the regional estimator is preferable in homogeneous and moderately heterogenous regions. The regional estimator is relatively more efficient in regions with a negative κ. The effect of intersite dependence on the normalised quantile estimator is well described by *Stedinger*'s [1983] formula, but the effect on the regional T-year event estimator is much less than this formula predicts. If the correlation structure is moderately heterogeneous, modest intersite dependence has only a small effect on the regional T-year event estimator. Thus, the PDS index-flood procedure is a robust and efficient estimation method.

The performance of the regional AMS and PDS index-flood methods has been compared with respect to the accuracy of T-year event estimators. For estimation in homogeneous regions, the PDS model is in general more efficient in regions with a negative κ, whereas the AMS model is preferable in regions with a positive κ. For estimation in heterogeneous regions, the PDS model is relatively more efficient, and for realistic degrees of heterogeneity it is superior to the AMS model for all κ. Thus, the PDS index-flood method is more robust than its AMS counterpart with respect to violation of the basic homogeneity assumption. The PDS index-flood model has also been compared to a modified AMS index-flood procedure in which only the skewness in the GEV distribution is estimated from regional data. This method has less strict assumptions with respect to regional homogeneity, but for realistic degrees of heterogeneity it is competitive only in regions with positive κ-values.

Procedures for grouping of sites into homogeneous regions and determination of regional parent distributions have been discussed. To illustrate these aspects, the AMS and PDS regional schemes were applied to flood records from New Zealand. For grouping of basins, a split-sample regionalisation approach based on catchment characteristics was adopted. To determine the optimal grouping of sites, a flood frequency variability measure based on L-moment ratios was employed. The application example revealed that the defined groups were more homogeneous, in terms of L-moment statistics, with respect to PDS than AMS data. A two-way grouping based on AAR was sufficient to attain homogeneity for PDS, whereas a further partitioning was necessary for AMS. To determine the regional parent distribution, L-moment ratio diagrams were constructed and a goodness-of-fit test was applied. The PDS data, in contrast to AMS data, were found to provide an unambiguous interpretation, supporting a GP distribution. Thus, for both identification of homogeneous regions and determination of regional distributions, the PDS approach has preferable properties.

To describe the variability of the regional PDS parameters and to include prior information about the index-flood parameter, an empirical Bayes estimation procedure has been introduced where the prior information is inferred from regional data using GLS regression. The regional mean model based on GLS regression has been shown to be a generalisation of the record-length-weighted average procedure usually adopted in index-flood modelling. The prior properties of the PDS parameters form the basis for quantile estimation at ungauged sites, and at gauged sites this information is combined with the site specific information. The empirical Bayes procedure provides an estimate of the T-year event as well as an estimate of the associated uncertainty. Two different empirical Bayes estimators have been introduced, respectively, a linear estimator that requires only the mean and the variance of the prior distributions to be specified and a parametric estimator where the families of prior distributions have to be defined.

The GLS and empirical Bayes estimation procedures have been applied to the New Zealand flood records. The application example clearly illustrated the importance of taking intersite dependence and regional heterogeneity into account when estimating the regional κ-parameter. In the case of a strongly heterogeneous correlation structure, the GLS procedure provides a more reasonable estimate of the regional κ-parameter than the record-length-weighted average procedure. Moreover, if intersite dependence is ignored, erroneous conclusions with respect to regional homogeneity may be drawn, and in such cases the uncertainty of the regional κ-estimator is seriously underestimated. The GLS procedure provides a reliable assessment of regional homogeneity and parameter uncertainty. A comparison of the two different Bayesian procedures revealed that the simple linear estimator is sufficient in most cases. Only when a more comprehensive Bayesian analysis is needed, the parametric estimator should be employed.

7. REFERENCES

Abild, J., Andersen, E.Y. and Rosbjerg, D., 1992, The climate of extreme winds at the Great Belt, Denmark, *J. Wind Eng. Industrial Aerodyn., 41*, 521-532.

Acreman, M.C. and Sinclair, C.D., 1986, Classification of drainage basins according to their physical characteristics: An application for flood frequency analysis in Scotland, *J. Hydrol., 84*, 365-380.

Acreman, M.C. and Wiltshire, S.E., 1989, The regions are dead; long live the regions. Methods of identifying and dispensing with regions for flood frequency analysis, FRIENDS in hydrology (Flow Regimes from International Experimental and Network Data Sets), *IAHS Publ. no. 187*, 175-188.

Ashkar, F. and Rousselle, J., 1987, Partial duration series modeling under the assumption of a Poissonian flood count, *J. Hydrol., 90*, 135-144.

Ashkar, F., Bobée, B. and Bernier, J., 1992, Separation of skewness: reality or regional artifact?, *J. Hydraul. Eng., ASCE, 118*(3), 460-475.

Bobée, B. and Rasmussen, P.F., 1995, Recent advances in flood frequency analysis, *Rev. Geophys., Suppl.*, 1111-1116.

Buishand, T.A., 1989, The partial duration series method with a fixed number of peaks, *J. Hydrol., 109*, 1-9.

Burn, D.H., 1989, Cluster analysis as applied to regional flood frequency, *J. Water Resour. Plann. Manage., ASCE, 115*(5), 567-582.

Burn, D.H., 1990a, An appraisal of the "region of influence" approach to flood frequency analysis, *Hydrol. Sci. J., 35*(2), 149-165.

Burn, D.H., 1990b, Evaluation of regional flood frequency analysis with a region of influence approach, *Water Resour. Res., 26*(10), 2257-2265.

Chowdhury, J.U., Stedinger, J.R. and Lu, L.-H., 1991, Goodness-of-fit tests for regional generalized extreme value flood distributions, *Water Resour. Res., 27*(7), 1765-1776.

Cunnane, C., 1973, A particular comparison of annual maxima and partial duration series methods of flood frequency prediction, *J. Hydrol., 18*, 257-271.

Cunnane, C., 1979, A note on the Poisson assumption in partial duration series models, *Water Resour. Res., 15*(2), 489-494.

Cunnane, C., 1987, Review of statistical models for flood frequency estimation, In: *Hydrologic Frequency Modeling* (ed. V.P. Singh), D. Reidel Publishing Company, 49-95.

Cunnane, C., 1988, Methods and merits of regional flood frequency analysis, *J. Hydrol., 100*, 269-290.

Dalrymple, T., 1960, Flood frequency analyses, *USGS, Water Supply Paper 1543A*, 11-51.

Davison, A.C. and Smith, R.L., 1990, Models for exceedances over high thresholds, *J. Royal Statist. Soc. B, 52*(3), 393-442.

Ekanayake, S.T. and Cruise, J.F., 1993, Comparisons of Weibull- and exponential-based partial duration stochastic flood models, *Stochastic Hydrol. Hydraul., 7*(4), 283-297.

Farquharson, F.A.K., Green, C.S., Meigh, J.R. and Sutcliffe, J.V., 1987, Comparison of flood frequency curves for many different regions of the world, In: *Regional Flood Frequency Analysis* (ed. V.P. Singh), D. Reidel Publishing Company, 223-256.

Fennessey, N. and Vogel, R.M., 1990, Regional flow-duration curves for ungauged sites in Massachusetts, *J. Water Resour. Plann. Manage., ASCE, 116*(4), 530-549.

Fitzgerald, D. L., 1989, Single station and regional analysis of daily rainfall extremes, *Stochastic Hydrol. Hydraul., 3*, 281-292.

Fuller, W.E., 1914, Flood flows, *Trans. Am. Soc. Civ. Eng., 77*, 564-617.

Gabriele, S. and Arnell, N., 1991, A hierarchical approach to regional flood frequency analysis, *Water Resour. Res., 27*(6), 1281-1289.

Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis, J.R., 1979, Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form, *Water Resour. Res., 15*(5), 1049-1054.

Greis, N.P. and Wood, E.F., 1981, Regional flood frequency estimation and network design, *Water Resour. Res., 17*(4), 1167-1177. Correction, *Water Resour. Res., 19*(2), 589-590, 1983.

Gustard, A., Roald, L.A., Demuth, S., Lumadjeng, H.S. and Gross, R., 1989, *Flow Regimes From Experimental and Network Data (FREND)*, Vol. 1, *Hydrological Studies*, UNESCO, IHP III, Project 6.1, Institute of Hydrology, Wallingford, UK.

Hebson, C.S. and Cunnane, C., 1987, Assessment of use of at-site and regional flood data for flood frequency estimation, In: *Hydrologic Frequency Modeling* (ed. V.P. Singh), D. Reidel Publishing Company, 433-448.

Hosking, J.R.M., 1988, The 4-parameter kappa distribution, *Res. Report RC13412*, IBM Research Division, Yorktown Heights, New York.

Hosking, J.R.M., 1990, L-moments: Analysis and estimation of distributions using linear combinations of order statistics, *J. Royal Statist. Soc. B, 52*(1), 105-124.

Hosking, J.R.M., 1991, Fortran routines for use with the method of L-moments, *Res. Report RC17097*, IBM Research Division, Yorktown Heights, New York.

Hosking, J.R.M. and Wallis, J.R., 1987, Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics*, *29*(3), 339-349.

Hosking, J.R.M. and Wallis, J.R., 1988, The effect of intersite dependence on regional flood frequency analysis, *Water Resour. Res.*, *24*(4), 588-600.

Hosking, J.R.M. and Wallis, J.R., 1993, Some statistics useful in regional frequency analysis, *Water Resour. Res.*, *29*(2), 271-281. Correction, *Water Resour. Res.*, *31*(1), 251, 1995.

Hosking, J.R.M. and Wallis, J.R., 1995, A comparison of unbiased and plotting-position estimators of L moments, *Water Resour. Res.*, *31*(8), 2019-2025.

Hosking, J.R.M., Wallis, J.R. and Wood, E.F., 1985a, An appraisal of the regional flood frequency procedure in the UK Flood Studies Report, *Hydrol. Sci. J.*, *30*(1), 85-109.

Hosking, J.R.M., Wallis, J.R. and Wood, E.F., 1985b, Estimation of the generalized extreme- value distribution by the method of probability-weighted moments, *Technometrics*, *27*(3), 251-261.

Houghton, J.C., 1978, Birth of a parent: The Wakeby distribution for modeling flood flows, *Water Resour. Res.*, *14*(6), 1105-1109.

Kirby, W.H. and Moss, M.E., 1987, Summary of flood-frequency analysis in the United States, In: *Analysis of Extraordinary Flood Events* (eds. W.H. Kirby, S.-Q. Hua and L.R. Beard), *J. Hydrol.*, *96*, 5-14.

Kuczera, G., 1982, Combining site-specific and regional information: An empirical Bayes approach, *Water Resour. Res.*, *18*(2), 306-314.

Kuczera, G., 1983, Effect of sampling uncertainty and spatial correlation on an empirical Bayes procedure for combining site and regional information, *J. Hydrol.*, *65*, 373-398.

Landwehr, J.M, Matalas, N.C. and Wallis, J.R., 1979, Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles, *Water Resour. Res.*, *15*(5), 1055-1064.

Lettenmaier, D.P., Wallis, J.R. and Wood, E.F., 1987, Effect of regional heterogeneity on flood frequency estimation, *Water Resour. Res.*, *23*(2), 313-323.

Lu, L.-H. and Stedinger, J.R., 1992a, Sampling variance of normalized GEV/ PWM quantile estimators and a regional homogeneity test, *J. Hydrol.*, *138*, 223-245.

Lu, L.-H. and Stedinger, J.R., 1992b, Variance of two- and three-parameter GEV/ PWM quantile estimators: formulae, confidence intervals, and a comparison, *J. Hydrol.*, *138*, 247-267.

Madsen, H. and Rosbjerg, D., 1994, Application of the partial duration series approach on a regional scale, Nordic Hydrological Conference, Tórshavn, *NHP-report no. 34*, 379-388.

Madsen, H. and Rosbjerg, D., 1995, On the modelling of extreme droughts, Modelling and Management of Sustainable Basin-scale Water Resources Systems (Proceedings of a Boulder Symposium, Juli 1995), *IAHS Publ. no. 231*, 377-385.

Madsen, H., Rosbjerg, D. and Harremoës, P., 1994, PDS-modelling and regional Bayesian estimation of extreme rainfalls, *Nordic Hydrol., 25*(4), 279-300.

Madsen, H., Rosbjerg, D. and Harremoës, P., 1995, Application of the Bayesian approach in regional analysis of extreme rainfalls, *Stochastic Hydrol. Hydraul., 9*(1), 77-88.

Matalas, N.C., Slack, J.R. and Wallis, J.R., 1975, Regional skew in search of a parent, *Water Resour. Res.*, *11*(6), 815-826.

Miquel, J., 1984, *Guide Pratique D'estimation des Probabilites de Crues*, Editions Eyrolles, Paris.

Naden, P.S., 1992, Analysis and use of peaks-over-threshold data in flood estimation, In: *Floods and Flood Management* (ed. A.J. Saul), Kluwer Academic Publishers, 131-143.

Nathan, R.J. and McMahon, T.A., 1990, Identification of homogeneous regions for the purposes of regionalisation, *J. Hydrol., 121*, 217-238.

NERC, 1975, *Flood Studies Report*, Vol I *Hydrological Studies*, Natural Environment Research Council, London.

Pearson, C.P., 1991a, New Zealand regional flood frequency analysis using L-moments, *J. Hydrol. NZ*, *30*(2), 53-64.

Pearson, C.P., 1991b, Regional flood frequency analysis for small New Zealand basins. 2. Flood frequency groups, *J. Hydrol. NZ, 30*(2), 77-92.

Pickands, J., 1975, Statistical inference using extreme order statistics, *Ann. Stat., 3*(1), 119-131.

Pilon, P.J. and Adamowski, K., 1992, The value of regional information to flood frequency analysis using the method of L-moments, *Can. J. Civ. Eng., 19*, 137-147.

Potter, K.W. and Lettenmaier, D.P., 1990, A comparison of regional flood frequency estimation methods using a resampling method, *Water Resour. Res., 26*(3), 415-424.

Rasmussen, P.F., 1991, The partial duration series approach to flood frequency analysis, *Series Paper, no. 55*, Inst. Hydrodyn. Hydraul. Eng., Tech. Univ. Denmark.

Rasmussen, P.F. and Rosbjerg, D., 1991, Application of Bayesian principles in regional flood frequency estimation, In: *Advances in Water Resources Technology* (ed. G. Tsakiris), Balkema, 65-75.

Rasmussen, P.F., Ashkar, F., Rosbjerg, D. and Bobée, B., 1994, The POT method for flood estimation: A review, In: *Stochastic and Statistical Methods in Hydrology and Environmental Engineering* (ed. K.W. Hipel), Kluwer Academic Publishers, Vol. 1, 15-26.

Rosbjerg, D., 1985, Estimation in partial duration series with independent and dependent peak values, *J. Hydrol., 76*, 183-195.

Rosbjerg, D., 1993, *Partial Duration Series in Water Resources*, Inst. Hydrodyn. Hydraul. Eng., Tech. Univ. Denmark.

Rosbjerg, D. and Knudsen, J., 1984, POT-estimation of extreme sea states and the benefit of using wind data, In: *Statistical Extremes and Applications* (ed. J. Tiago de Oliveira), D. Reidel Publishing Company, 611-620.

Rosbjerg, D. and Madsen, H., 1992, On the choice of threshold level in partial duration series, Nordic Hydrological Conference, Alta, *NHP-report no. 30*, 604-615.

Rosbjerg, D. and Madsen, H., 1996, The role of regional information in estimation of extreme point rainfalls, *Atmos. Res., 42*, 113-122.

Rosbjerg, D., Madsen, H. and Rasmussen, P.F., 1992, Prediction in partial duration series with generalized Pareto-distributed exceedances, *Water Resour. Res., 28*(11), 3001-3010.

Rosbjerg, D., Rasmussen, P.F. and Madsen, H., 1991, Modelling of exceedances in partial duration series, *Proceedings of the International Hydrology and Water Resources Symposium*, Perth, 755-760.

Rousselle, J. and Hindie, F., 1976, Incertitude dans les debits de crues: Approche Bayesienne, *J. Hydrol., 30*, 341-349.

Shane, R.M. and Lynn, W.R., 1964, Mathematical model for flood risk evaluation, *J. Hydraul. Div. ASCE, 90(HY6)*, 1-20.

Smith, C.D., 1993, Selection of design flood frequencies for hydraulic structures, *Can. J. Civ. Eng., 20*, 582-586.

Smith, J.A., 1989, Regional flood frequency analysis using extreme order statistics of the annual peak record, *Water Resour. Res., 25*(2), 311-317.

Stedinger, J.R., 1983, Estimating a regional flood frequency distribution, *Water Resour. Res., 19*(2), 503-510.

Stedinger, J.R. and Lu, L.-H., 1995, Appraisal of regional and index flood quantile estimators, *Stochastic Hydrol. Hydraul., 9*(1), 49-75.

Stedinger, J.R. and Tasker, G.D., 1985, Regional hydrologic analysis, 1. Ordinary, weighted and generalized least squares compared, *Water Resour. Res., 21*(9), 1421-1432. Correction, *Water Resour. Res., 22*(5), 844, 1986.

Stedinger, J.R. and Tasker, G.D., 1986, Regional hydrologic analysis, 2. Model-error estimators, estimation of sigma and log-Pearson Type 3 distributions, *Water Resour. Res., 22*(10), 1487-1499.

Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E., 1993, Frequency analysis of extreme events, *Handbook of Hydrology* (ed. D.R. Maidment), Chapter 18, McGraw-Hill.

Tallaksen, L.M., Madsen, H. and Clausen, B., 1997, On the definition and modelling of streamflow drought duration and deficit volume, *Hydrol. Sci. J., 42*(1), 15-33.

Tasker, G.D., 1980, Hydrologic regression with weighted least squares, *Water Resour. Res., 16*(6), 1107-1113.

Todorovic, P. and Zelenhasic, E., 1970, A stochastic model for flood analysis, *Water Resour. Res., 6*(6), 1641-1648.

USWRC, 1982, United States Water Resources Council: Guidelines for determining flood flow frequency, *Bull. 17B*, Hydrol. Comm., Water Resour. Counc., Washington, D.C.

Van Montfort, M.A.J. and Witter, J.V., 1986, The generalized Pareto distribution applied to rainfall depths, *Hydrol. Sci. J., 31*(2), 151-162.

Vogel, R.M. and Fennessay, N.M., 1993, L moment diagrams should replace product moment diagrams, *Water Resour. Res., 29*(6), 1745-1752.

Vogel, R.M. and Wilson, I., 1996, Probability Distribution of Annual Maximum, Mean, and Minimum Streamflows in the United States, *J. Hydrol. Eng., ASCE, 1*(2), 69-76.

Vogel, R.M., McMahon, T.A. and Chiew, F.H.S., 1993a, Floodflow frequency model selection in Australia, *J. Hydrol., 146*, 421-449.

Vogel, R.M., Thomas, W.O. and McMahon, T.A., 1993b, Flood-flow frequency model selection in Southwestern United States, *J. Water Resour. Plann. Manage., ASCE, 119*(3), 353-366.

Wallis, J.R., 1980, Risk and uncertainties in the evaluation of flood events for the design of hydraulic structures, In: *Piene e Siccità* (eds. E. Guggino, G. Rossi, and E. Todini), Fondazione Politecnica del Mediterraneo, Catania, Italy, 3-36.

Wallis, J.R. and Wood, E.F., 1985, Relative accuracy of log Pearson III procedures, *J. Hydraul. Eng., ASCE, 111*(7), 1043-1056.

Wang, Q.J., 1991, The POT model described by the generalized Pareto distribution with Poisson arrival rate, *J. Hydrol., 129*, 263-280.

Wiltshire, S.E., 1985, Grouping basins for regional flood frequency analysis, *Hydrol. Sci. J., 30*(1), 151-159.

Wiltshire, S.E., 1986a, Regional flood frequency analysis I: Homogeneity statistics, *Hydrol. Sci. J., 31*(3), 321-333.

Wiltshire, S.E., 1986b, Regional flood frequency analysis II: Multivariate classification of drainage basins in Britain, *Hydrol. Sci. J., 31*(3), 335-346.

Wiltshire, S.E., 1986c, Identification of homogeneous regions for flood frequency analysis, *J. Hydrol., 84*, 287-302.

WMO (World Meteorological Organization), 1989, Statistical distributions for flood frequency analysis, Operational Hydrology, *Rep. no. 33, WMO Publ. no. 718*, Geneva, Switzerland.

Zelenhasic, E., 1970, Theoretical probability distributions for flood peaks, *Hydrology Paper, no. 42*, Colorado State University, Fort Collins, Colorado.

Zelenhasic, E. and Salvai, A., 1987, A method of streamflow drought analysis, *Water Resour. Res., 23*(1), 156-168.

Zrinji, Z. and Burn, D.H., 1994, Flood frequency analysis for ungauged sites using a region of influence approach, *J. Hydrol., 153*, 1-21.